

## DATA ANALYSIS FOR EFFECTIVE MONITORING OF PARTIALLY SHADED PHOTOVOLTAIC SYSTEMS.

Odysseas Tsafarakis<sup>1</sup>, Kostas Sinapis<sup>2</sup>, Wilfried G.J.H.M. van Sark<sup>3</sup>

<sup>1,3</sup>Utrecht University, Copernicus institute of Sustainable Development, Heidelberglaan 2, 3584 CS, Utrecht, The Netherlands

<sup>2</sup>Solar Energy Application Centre, HTC 21, 5656 AE, Eindhoven, The Netherlands

<sup>1,2,3</sup>E: O.Tsafarakis@uu.nl, sinapis@ecn.nl, W.G.J.H.M.vanSark@uu.nl

**ABSTRACT:** The purpose of this study is the development of an algorithmic tool to automate the process of analyzing monitoring data of partially shaded PV systems. The approach is to compare long-term and high resolution yield data of a partially shaded Photovoltaic (PV) system (investigated PV) with the yield data of an unshaded PV system or the tilted irradiance (reference PV system) and automatically detect the energy loss due to the expected shadow, caused by any surrounding obstacles and distinct it from any additional energy loss due to other malfunctions.

**Keywords:** see enclosed list of keywords

### I. INTRODUCTION

The majority of the installed PV systems in the Netherlands are relatively small-scale installations, with an average capacity of 3.5 kWp[1]. About 70% of them is placed on residential rooftops[2], where different objects might obstruct the irradiance reaching the PV module and affects its energy output. Such small-scale, low-cost PV installations are mostly monitored through relatively simple, low-cost, data loggers of AC power, without any further equipment for solar measurements, i.e., pyranometers.

The ultimate aim of our research is to find an effective manner that allows monitoring large numbers of such installations in an automated manner. The proposed method offers the possibility to monitor partly shaded systems by excluding the effects of constant and unavoidable shadow from the performance evaluation and to detect faults that might occur during their operation.

### II SCIENTIFIC INNOVATION

The innovation of the proposed method is its independence from the use of irradiance data. The performance evaluation is not based on the estimation of the energy generation/production of the studied PV systems and simulation of the losses due to shading (with the use of irradiance data and models) and its comparison with the real production[3]–[5] or the comparison of the production with the tilted irradiance, through the calculation of the performance ratio[5], [6]. It is based on the performance comparison and the detection of repeated or non-repeated errors/malfunctions between neighboring PV systems with similar characteristics (tilt and orientation).

### III. METHODOLOGY

The basic principle of the proposed method/algorithm, is to study the daily persistence of any occurred error between the investigated PV and the reference PV system and identify these daily repeated errors, characterize them as shadows and create the “**shadow profile**” of the studied PV, i.e. a table with the starting and ending dates of the shadow(s) during the

year and the expected starting and ending times of the shadow(s) during each day. Any difference between the yield of the investigated PV and the reference PV system, outside of the limits of these shadow profile time intervals is characterized as a malfunction and within these limits as shadow. An example of the shadow profile is presented in table 1. In this case, 5 different shadow periods are forming the shadow profile of the studied PV system.

During the analysis, the fluctuation of the Mean Absolute Error (MAE) between the production of the systems is studied ( $P_{\text{reference}} - P_{\text{shaded}}$ ). Every timestamp with MAE higher than the normal is characterized as error and it is studied, in order to be characterized as shadow or malfunction. In order to set a threshold between the normal and the not normal operation, the iterative method Ran.Sa.C. (Random Sample Consensus) [7] is used.

Due to the nature of earth’s orbit around the sun, the time interval of each shadow effect is varying from day to day. Additionally, due to the nature of the diffuse irradiance, the shadow is not present on cloudy days, where 5 days in row without any effect of shadow could appear. These facts are making the shading detection more difficult than a simple pattern recognition problem since different groups of hours and dates are forming the shadow profile. The clustering method D.B.S.C.A.N.[8] (Density-based Spatial Clustering of Applications with Noise) is used in order to identify shadows during different periods.

Thereafter the continuity of all the shaded periods of the “shadow profile” table is studied and presented in a “Date vs Time” plot, which demonstrates the change of the shadow effect during a long time studied period. We denote this as the “**Shadow story**”. An example of the shadow story is presented in figure 1.

### IV. RESULTS

As mentioned above, an example a partially shaded PV system is compared with a non-shaded PV system and their “**shadow window**”, i.e. the shadow periods of the system during the year (table I) and its **Shadow Story** (figure 1), i.e. the day vs time plot, are presented. The green marks in figure 1 indicate

the measurements without a presence of error between the studied (partially shaded) and the reference (non-shaded) PV system. A shading obstacle was placed from August 10 to November 11 and impacts the PV system's performance during the morning hours. Since this error is repeated for long period, it is identified by the algorithm as shadow (black – higher than threshold difference – and blue – lower than threshold difference – marks) and not as malfunction. The shadow from the neighboring rooftop also affects the system in the late afternoon is also recognized as shadow.

N.	Start day	End day	Start time (HH:MM)	End time (HH:MM)	Cause
1	Aug 8	Oct 13	9:05	12:05	Pole
2	Oct 18	Nov 11	9:15	13:55	Pole
3	Jun 1	Jun 29	18:40	19:40	Roof
4	Jul 16	Aug 22	18:50	19:40	Roof
5	Aug 28	Sep 20	19:25	19:40	Roof

**Table I:** shadow windows example, where every row is a different shadow

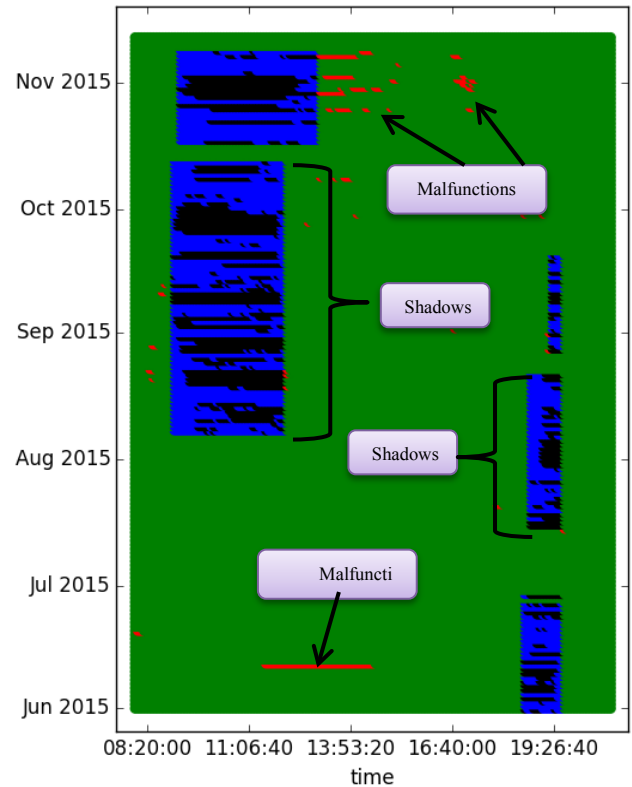
Major remarks on the plot:

- At June 12 for external reasons, an obstacle was placed in front of the system, from 11:35 to 14:30. It is clearly within the hours where the pole is affecting the system (when it is placed). However, the algorithm recognized this incident as malfunction, since it is the only performance reduction, which occurred in that period and it is not in the table 1. A similar incident (roof maintenance) occurred for some days between October 26 and November 6. These dates are in shadow profile (shadow 2 in table I), however bigger part of the error was outside of the time intervals of the shadow profile, thus they marked as red.
- As depicted in figure 1, the shadow is shifting from August to November and impacts the system in later hours, due to the fact that the day is becoming shorter. In this case the role of the date clustering through DBSCAN is important, since the shadow period is separated in two smaller periods (shadows 1 & 2, Table I) and each one has different start and end time.

## V. FURTHER DISCUSSION OF THE RESULTS

As depicted in figure 1, within the time intervals of the shadow profile (Blue points) there is a large numbers of measurements without error. It can be assumed that the reason of this fact is the high diffuse irradiance, which, in contrast with the direct, doesn't obstructed by the shade causing object. This assumption is proved in this paragraph, where the blue points of figure I are further studied.

In Figure II, the plot of figure I is replotted with one major difference. The blue points of figure I are further studied and colored depending on the solar irradiance and specifically on



**Figure 1:** Shadow Story

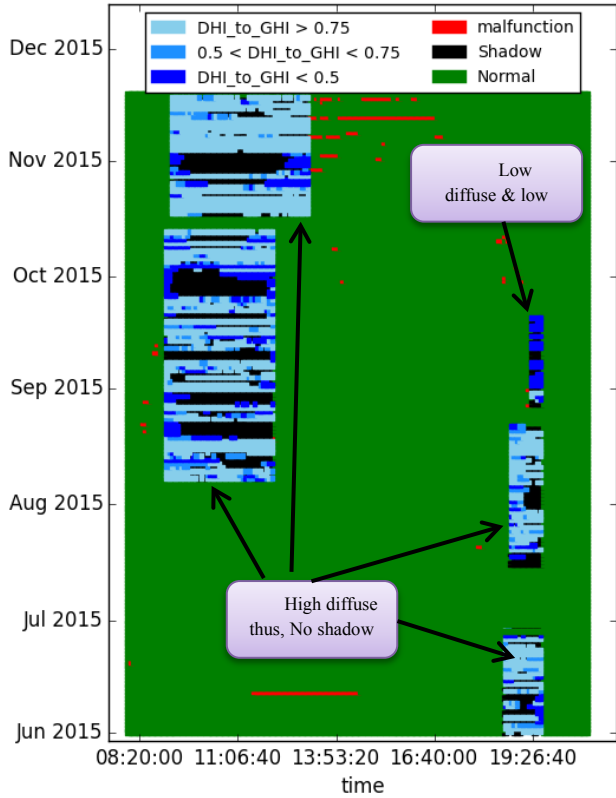
the ratio of the diffuse horizontal irradiance (DHI) to the global horizontal irradiance (GHI) or  $R_{DHI}$ .

As depicted in figure II, in the majority of the points,  $R_{DHI}$  is higher than 0.5 (lighter blues) i.e. the DHI is higher than the direct normal irradiance (DNI). Further analysis of the blue points is presented in the scatterplot of figure III. In this scatter plot the three red lines are the 3 quantiles and the green line is the percentile where the  $R_{DHI}$  is 0.5. As it is clear in figure III, the 87% of the sample is above 0.5 thus the DHI is higher than the DNI, the irradiance which causes the shadow.

However it is very interesting that the upper right shadow (September, 18:00 to 19:00) of the plots I & II has still only dark blue dots in figure II. This fact is explained in figure IV, where the blue points are studied again and recolored. In this figure the blue dots are colored depending on the global tilted irradiance (GTI) and there are colored yellow if  $GTI < 100 \text{ W/m}^2$ . Under this filter, all the blue marks of the specific profile are colored yellow, thus the irradiance is very low, in order to create any shadow effects.

## VI. SUMMARY OF THE WORK

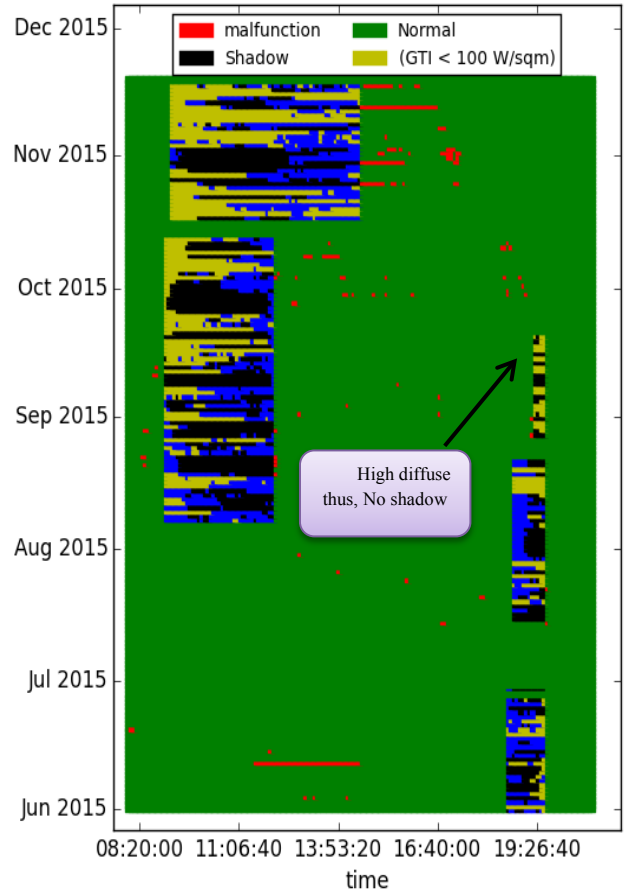
The proposed method is offering a monitoring solution to small residential PV systems, without the use of high-cost monitoring equipment which is completely cost ineffective



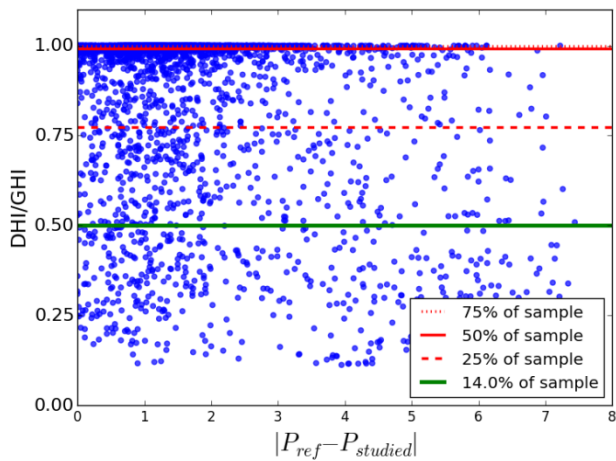
**Figure II:** Shadow story with measurements without fault inside shadow profile colored depending on the ratio DHI/GHI

for this kind of installations. It only uses the PAC of two neighboring PV systems, with similar tilt and orientation and not necessary similar capacity (since the production can be normalized by divided with the  $W_{peak}$  of the system). The developed algorithm automatically isolates the timestamps with the non-normal operation and studies their concentration during the hours of the day and their

continuity during the studied period and provides the owner of the system with the shadow story, a plot which demonstrates the effect of the shadow on his PV system during the year



**Figure IV:** Shadow story with unshaded measurements inside shadow profile colored depending on the GTI



**Figure III** – Further analysis of the non fault measurements of shadow profiles

## VII. REFERENCES

- [1] W. van Sark, S. Hart, M. de Jong, P. de Rijk, P. Moraitis, B. B. Kausika, and H. van der Velde, “‘Counting the Sun’ – a Dutch Public Awareness Campaign on Pv Performance,” *29th Eur. Photovolt. Sol. Energy Conf. Exhib.*, vol. 2014, pp. 3545–3548, 2014.
- [2] Centraal Bureau voor de Statistiek (CBS), “Bijgeplaatst vermogen zonnestroom bijgesteld.” [Online]. Available: <https://www.cbs.nl/nl-nl/achtergrond/2015/51/bijgeplaatst-vermogen-zonnestroom-bijgesteld>.
- [3] A. Drews, A. C. de Keizer, H. G. Beyer, E. Lorenz, J. Betcke, W. G. J. H. M. van Sark, W. Heydenreich, E. Wiemken, S. Stettler, P. Toggweiler, S. Bofinger, M. Schneider, G. Heilscher, and D. Heinemann, “Monitoring and remote failure detection of grid-connected PV systems based on satellite observations,” *Sol. Energy*, vol. 81, no. 4, pp. 548–564, 2007.
- [4] S. Silvestre, A. Chouder, and E. Karatepe, “Automatic fault detection in grid connected PV systems,” vol. 94, pp. 119–127, 2013.
- [5] A. Woyte, M. Richter, D. Moser, M. Green, S. Mau, and H. G. Beyer, *Analytical Monitoring of Grid-connected Photovoltaic Systems*, vol. 13, no. 2. 2014.
- [6] R. Eke and A. Senturk, “Monitoring the performance of single and triple junction amorphous silicon modules in two building integrated photovoltaic (BIPV) installations,” *Appl. Energy*, vol. 109, pp. 154–162, 2013.
- [7] M. a Fischler and R. C. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [8] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” *Proc. 2nd Int. Conf. Knowl. Discov. Data Min.*, pp. 226–231, 1996.